

Pricing of Drugs with Heterogeneous Health Insurance Coverage

Ida Ferrara*

Department of Economics
York University

Paul Missios

Department of Economics
Ryerson University

January 19, 2010

Abstract

In this paper, we examine the role of insurance coverage in explaining the generic competition paradox in a two-stage game involving a single producer of brand-name drugs and n quantity-competing producers of generic drugs. Independently of brand loyalty, which some studies rely upon to explain the paradox, we show that heterogeneity in insurance coverage may result in higher prices of brand-name drugs following generic entry. With market segmentation based on insurance coverage present in both the pre- and post-entry stages, the paradox can arise when the two types of drugs are highly substitutable and the market is quite profitable but does not have to arise when the two types of drugs are highly differentiated. However, with market segmentation occurring only after generic entry, the paradox can arise when the two types of drugs are weakly substitutables, provided, however, that the industry is not very profitable. In both cases, that is, when market segmentation is present in the pre-entry stage and when it is not, the paradox becomes more likely to arise as the market expands and/or insurance companies decrease deductibles applied on the purchase of generic drugs.

JEL Classifications: L11, L13, I12

Keywords: brand-name pricing, generic entry, generic competition paradox, health insurance.

*Corresponding author: Department of Economics, York University, 4700 Keele Street, Toronto, ON, Canada, M3J 1P3, e-mail: iferrara@yorku.ca.

1 Introduction

The development and growth of the generic pharmaceutical industry over the past 25 years has come in response to rising healthcare costs. In 2004, healthcare costs represented 15.3 percent of GDP in the United States, the highest share among OECD countries, followed by Switzerland (11.6 percent), Germany (10.9 percent), and France (10.5 percent), all above the OECD average of 8.9 percent (OECD, 2007). Private expenditures per capita were also the highest in the United States, more than double the private expenditures per capita of any other OECD country. For all OECD countries, the rate of growth of health spending per capita increased over the period 1999-2004 by more than 5 percent per year. With pharmaceutical expenditures representing 10 to 25 percent of health expenditures (OECD), many countries have attempted to promote the use of generic drugs in a variety of ways in order to keep healthcare costs down while maintaining or even increasing accessibility to pharmaceuticals and retaining incentives to invest in innovation and research and development.

In 1984, the Drug Price Competition and Patent Term Restoration Act, also known as the Waxman-Hatch Act, was introduced in the U.S. in order to improve generic competition by lowering barriers to entry for generic drugs and to increase patent terms for new drugs delayed by complicated and time-consuming approval procedures of the U.S. Food and Drug Administration (FDA), the agency responsible for the safety and efficacy of drugs. Under this legislation, (duplicative) testing for generic drugs was eliminated and replaced by the requirement that an Abbreviated New Drug Application (ANDA) be submitted by generic entrants demonstrating the equivalence between their products and the original (brand-name) drugs. Not surprisingly, the entry of generic drugs into the pharmaceutical market intensified dramatically following the introduction of the Waxman-Hatch Act, and as a result of the expiration of patents on many high-sales-volume brand-name drugs (Frank and Salkever, 1997). In response to the increased market share of generic drugs and lower prices of pharmaceuticals overall, the prices of brand-name drugs did not fall consistently with the predictions of traditional market entry models; instead, they were often observed to increase. This phenomenon is often referred to as the generic competition paradox, or GCP (Scherer, 1993).

The first evidence of the paradox was presented by Wagner and Duffy (1988) who found substantial price increases associated with entry despite significant decreases in generic prices among top selling name brand drugs. Several other instances of support for the paradox were provided in subsequent studies, including those by Grabowski and Vernon (1992), Frank and Salkever (1997), Perloff et al. (1995), and, to a lesser extent, Caves et al. (1991). Grabowski and Vernon, using data on 18 major drugs in the mid-1980s, showed that price increases for name brands increased on average by 7 percent after entry and continued to increase in the following year. Frank and Salkever utilized data on the patent expiration of 45 drugs facing competition for the first time between 1984 and 1987, and found that, although significant market share was lost upon entry by brand-name producers, the price of their product generally increased. Similarly, Perloff et al. showed brand-name price increases using data from the mid-1980s from the U.S. anti-ulcer drug market. Although Caves et al. did not observe price increases in their study of 30 drugs between 1976 and 1987, they found prices to decrease by only small amounts following entry (2 percent on patent loss, although this loss increased with the number of entrants) and by far less than the price decrease experienced by the entrants. In other studies, including the work by Wiggins and Maness (1998) on 98 anti-infectives from 1984 to 1990, no evidence in support of the paradox was detected.

Traditional oligopolistic models of entry suggest that the increased competition caused by generic entry should drive prices down for all firms, as illustrated by a move from monopoly to duopoly with homogeneous goods. While most, if not all, models attempting to explain the paradox would suggest that *average* prices fall after entry, the standard models do not explain why the price charged by the incumbent firm could increase after entry. Several explanations have been proposed to theoretically support the empirical finding that prices of brand-name drugs increase after entry. For example, but outside of the realm of the pharmaceutical industry, models of entry-induced price increases in oligopolistic or monopolistically competitive markets, including those by Satterthwaite (1979), Salop (1979), Rosenthal (1990), suggest that economies of scale or specific demand curve changes can lead to post-entry price increases.¹ Due to the nature of pharmaceutical production, economies of

¹Davis, Murphy and Topel (2004) suggest that differentiation and segmentation may lead to price increases after entry beyond drugs to other products such as Microsoft's Windows.

scale are not typically present and therefore this is not a likely explanation for the paradox. However, several papers employ changes in the elasticity of demand to explain the paradox, including the brand-loyalty models of Caves et al. (1991), Grabowski and Vernon (1992), Frank and Salkever (1997), and Kamien and Zang (1999).² In these models, *exogenous* segmentation of the market occurs upon entry, as one group of consumers is price sensitive while another is not.³ The segmentation is exogenous in the sense that the individual groups exist separately in the market prior to entry but the brand name producer is not permitted to choose whether or not they want to serve only one group prior to entry, and the size of each group remains constant before and after entry.

Other studies attempting to explain the paradox introduce price stickiness into models due to imperfectly informed doctors (Bhattacharya and Vogt, 2003), product differentiation with collusion or price competition (Perloff et al, 1995), or quality differences (Berndt et al., 1993; Griliches and Cockbrun, 1994). Bhattacharya and Vogt argue that doctors' stock of knowledge about the presence and efficacy of new generics evolves slowly and is manipulated by producers through advertising. Perloff et al. show that the paradox can occur when products are significantly differentiated in product space. Pre-entry, the firm lowers its price to serve segments of the market located far away from its product in its characteristics, but increases its price once entry occurs and those consumers switch to the entrant's product. Product location is exogenous, and the paradox is not possible for cases in which the entrant's product and the incumbent's product are closely related (little product differentiation). Berndt et al. and Griliches and Cockbrun have that price increases are generated by quality improvements, and that prices increase over the life of a product (although increase more slowly post-entry).

In this paper, we combine some of the features of the models of previous studies, including product differentiation and brand loyalty, but focus on the role of insurance coverage in the segmentation of the market. Specifically, we examine the endogenous segmentation of the market by the brand-name producer, both before and after entry, to determine whether or

²Kamien and Zang focus on the introduction of generics by brand-name firms prior to entry rather than on the segmentation by brand loyalty itself.

³Exogenous segmentation by physicians based on insurance coverage is recently considered by Ferrara and Kong (2008).

not insurance coverage can explain the paradox, and its relation to the other theories of the literature.

The relevance of insurance in pricing decisions has empirical support. Hellerstein (1994), for example, using prescription data from the eight largest therapeutic drug classes, describes how most individual doctors prescribe both brand-name and generic drugs (suggesting that a lack of awareness or knowledge is not driving the prescription decision). Furthermore, doctors with higher fractions of Medicaid, Medicare, HMO, and privately insured patients are more likely to prescribe generics, although the links are not particularly strong (or even negative) for certain drug classes. Pavcnik (2002) suggests that brand-name pricing is very sensitive to out-of-pocket expenditures, and estimates, using data from Germany, that the price adjustment to an exogenous change in insurance coverage ranges between 10 percent and 26 percent. In a cross-country study, Danzon and Chao (2000) show that the effect of generic competition on brand-name prices depends on the insurance coverage and pricing regime, and conclude that countries with fee pricing (like the U.S.) tend to experience large decreases but countries with strict reimbursement regulation and insurance (like France, Italy, and Japan) tend to experience price increases.⁴

In the present analysis, we construct a model in which consumers differ on the basis of coverage, and the brand-name producer can choose, through its price for the brand-name drug and taking into account the impact of its own decisions on generic pricing, which consumers it wants to target and which consumers it wants to leave out of the market (pre-entry) or to the generic producers (post-entry). The inclusion of a parameter, q , into the utility function describing consumers' preferences, which captures the perceived quality differential between brand-name and generic drugs, allows for a separation between the price effects of brand loyalty as reflected in q and the price effects of segmentation induced by insurance coverage heterogeneity. We thus derive conditions under which the generic competition paradox occurs in instances in which brand loyalty (or the perceived quality differential) alone does not give rise to price increases for brand-name drugs when generic drugs are introduced so that the paradox can only be attributed to insurance coverage

⁴Other papers, such as that by Anis (1992), focus on the effectiveness of a reimbursement regime as a solution to the principal-agent problem in doctors' prescription decisions.

considerations. In the context of the model, we revisit the brand loyalty argument to explain the paradox and show that brand loyalty does not always yield increases in the price of brand-name drugs after generic entry; whether brand loyalty leads to price increases depends upon other parameters of the model such as the degree of product differentiation, production costs, and the willingness to pay for pharmaceuticals.

When the brand loyalty argument does not apply and the market is segmented on the basis of insurance coverage (at least after the introduction of generic drugs), we show that the price of brand-name drugs can increase following generic entry at low levels of product differentiation and can decrease at high levels of product differentiation. The extent of substitutability between brand-name and generic drugs is not by itself a key factor in determining whether the GCP arises and does not necessarily involve unambiguous effects on the likelihood of the paradox, as in Perloff et al. (1995); instead, a combination of product differentiation and market profitability, along with market size and preferential treatment of purchases of generic drugs by insurance companies, is a stronger determinant of when the paradox is likely to emerge so that lower product differentiation can support the paradox and higher product differentiation can yield no paradox.

When brand-name and generic drugs are close substitutes, the producer of brand-name drugs responds to generic entry by supplying to fewer consumers (those who have better insurance coverage) and can thus charge a higher price; such a strategy becomes less appealing as the market becomes less profitable so that the GCP is more likely to occur when the two types of drugs are not very differentiated but the market is very profitable. What happens if brand-name and generic drugs are not close substitutes depends on whether segmentation occurs prior to generic entry: if it does, the producer of brand-name drugs responds to generic entry by supplying to more consumers and must thus charge a lower price so that the paradox does not occur when the two types of drugs are highly differentiated and is even less likely as the market becomes more profitable; if it does not, it responds to generic entry by supplying to fewer consumers (moving away from providing drugs to the entire market) especially when the market is less profitable so that the GCP is more likely to occur at high levels of product differentiation but low levels of market profitability. The larger the market and/or the lower the deductible insurance companies apply on the purchase of generic drugs

relative to the deductible on the purchase of brand-name drugs, the stronger the incentive to segment is and thus the more likely the GCP arises.

We structure the remainder of the paper as follows: in section 2, we introduce the model and derive the equilibrium derived before and after generic entry both in the absence of health insurance coverage considerations (that is, consumers are homogenous) in sub-section 2.1 and when consumers differ in their health insurance coverage in sub-section 2.2; in section 3, we provide a numerical example; in section 4, we give concluding remarks.

2 The Model

In this paper, the pharmaceutical market is characterized by two products (the brand-name drug, produced by a monopolist, and its generic substitute, produced by n quantity-competing firms) and consumers differing in their insurance coverage as captured by θ , which is uniformly distributed over the interval $[\underline{\theta}, 1]$, with $\underline{\theta} > 0$. Specifically, θ denotes the fraction of expenditures on drugs a consumer pays out of his/her pocket with $(1 - \theta)$ thus reflecting the covered portion; hence, if θ is equal to 1, the consumer has no coverage. As the lower bound of θ is $\underline{\theta}$, 100 percent coverage is not available; in other words, a co-payment or deductible is always present and is defined as a fraction of total expenditures as opposed to some fixed amount that is independent of the level of spending. To account for a differential co-payment or deductible system which favours the purchase of generic drugs,⁵ the parameter t is introduced, with $t \in (0, 1)$, to capture the reduction in deductible or co-payment a consumer is entitled to if he/she buys generic drugs as opposed to brand-name drugs.⁶ Social insurance systems in France, Italy, Germany, Japan, and the United Kingdom, and

⁵Although the practice of reference-based pricing (the ability of a drug coverage plan to only reimburse a patient up to the amount of the lowest cost comparable drug available in the market) is quite well documented for public drug coverage plans, its existence is rather obscure among private providers of drug coverage plans. Most private insurance providers in the U.S. do in fact serve patients through their employers. As drug coverage benefits can be an important premise for an employee when considering which company to work for, employers would need to be competitive; the expectation is then that the private drug insurance providers would also need to be competitive and ensure its clients that they can offer a better coverage at a reasonable cost. Better coverage can naturally be interpreted as coverage for brand-name drugs as well as generic drugs. We do not consider alternate forms of regulation, such as price cap measures. For an empirical study of the effects of price caps and reference pricing on off-patent drugs, see Brekke et al. (2009).

⁶The consumer with $\theta = 1$ has no coverage for brand-name drugs but a coverage equal to t per dollar spent on generic drugs.

private insurance policies in Canada and the U.S. typically include co-payments with levels dependent upon whether the brand-name or generic good is purchased (Danzon and Chao, 2000). For example, most private insurance plans in the U.S. (83 percent in 2003, up from 25 percent in 1993) include a higher reimbursement for generic drugs (Dietz, 2004), and the brand-name co-payment exceeds the generic co-payment in the social insurance system in Germany.⁷

As in Caminal and Vives (1999), the utility function of a representative consumer is given by

$$U(x_0, x_b, x_g) = x_0 + [\alpha + (\beta - \gamma)q]x_b + [\alpha - (\beta - \gamma)q]x_g - \frac{1}{2}(\beta x_b^2 + 2\gamma x_b x_g + \beta x_g^2), \quad (1)$$

where x_0 represents units of a competitively produced numeraire good, x_b and x_g denote the consumption levels of brand-name and generic drugs, and α , β , γ , and q are positive parameters with $\beta > \gamma$, $\alpha > (\beta - \gamma)q$, and $q \geq 1$. Specifically, the ratio γ/β represents the degree of product differentiation so that the two products are perfect substitutes if $\gamma/\beta = 1$ and independent if $\gamma/\beta = 0$. While it may be that the brand-name and the generic types of drugs constitute the same product, there are several options available to producers to differentiate them. For example, producers may differentiate on the basis of taste, delivery system, oral dosage forms (capsules, coated pills, etc.), unit dosage packaging, and pill shape (Hurwitz and Caves, 1988); hence, we rule out the possibility that the two types of drugs are perfect substitutes, that is, $\beta > \gamma$ or $\gamma/\beta < 1$. The parameter q represents instead the quality differential between brand-name and generic drugs as perceived by consumers so that an increase in q triggers an increase in the marginal willingness to pay for brand-name drugs (i.e., $\alpha + (\beta - \gamma)q$) but a decrease in the marginal willingness to pay for generic drugs (i.e., $\alpha - (\beta - \gamma)q$). The parameter q can be interpreted as a measure of (or proxy for) brand loyalty since consumers' attachment to a particular brand typically stems from a perception of quality superiority. In essence, q captures the presence and the extent of imperfect information about the quality of generic drugs so that, even in instances in which the two types of drugs have very similar characteristics (that is, γ/β is close to unity),

⁷ "Tiered" systems, with different prices for generic formulary drugs, brand-name formulary drugs, and non-formulary drugs, have also been increasing in popularity over time (Dietz, 2004).

consumers may be reluctant to switching to the less expensive type when it becomes available, thus remaining loyal to the brand-name type, if they perceive it to be of inferior quality. As it is shown later, the inclusion of q into the utility function allows for a separation between the price effects of brand loyalty and the price effects of insurance coverage heterogeneity.

The budget constraint of type- θ consumer is given by

$$m = x_0 + \theta p_b x_b + (1 - t) \theta p_g x_g, \quad (2)$$

where m is income, p_b and p_g are the price levels of the brand-name and generic drugs, θ is the insurance factor or a parameter that captures the amount of insurance coverage for brand-name drugs, and t is the additional coverage on the purchase of generic drugs. A type- θ consumer thus pays θ for each dollar spent on brand-name drugs but $\theta(1 - t)$ for each dollar spent on generic drugs, with $0 < t < 1$, so that lower values of θ indicate greater insurance coverage. Income is then divided between spending on the numeraire good, uncovered spending on brand-name drugs, and (if available) uncovered spending on generic drugs.

Upon substitution for x_0 from (2) into (1), utility maximization with respect to x_b and x_g yields

$$x_b = \frac{1}{\beta} \{[\alpha + (\beta - \gamma) q] - \theta p_b - \gamma x_g\} \quad (3)$$

and

$$x_g = \frac{1}{\beta} \{[\alpha - (\beta - \gamma) q] - (1 - t) \theta p_g - \gamma x_b\}, \quad (4)$$

which can be re-written as

$$x_b = \frac{\alpha + (\beta + \gamma) q}{\beta + \gamma} - \left(\frac{1}{\beta^2 - \gamma^2} \right) [\beta \theta p_b - \gamma (1 - t) \theta p_g] \quad (5)$$

and

$$x_g = \frac{\alpha - (\beta + \gamma) q}{\beta + \gamma} - \left(\frac{1}{\beta^2 - \gamma^2} \right) [\beta (1 - t) \theta p_g - \gamma \theta p_b], \quad (6)$$

where $(\beta^2 - \gamma^2) > 0$ by the second-order conditions.

In deriving market demand functions, three cases are considered:

- CASE A (pre-entry, no segmentation): generic drugs are not available and brand-name drugs are supplied to the entire continuum of consumers, that is, $1 - \underline{\theta}$;

- CASE B (pre-entry, segmentation): generic drugs are not available and brand-name drugs are supplied to a segment of the continuum, that is, $\theta_B - \underline{\theta}$ where $\underline{\theta} < \theta_B < 1$ with θ_B endogenously chosen;⁸
- CASE C (post-entry, segmentation): generic drugs are available to the entire continuum of consumers and brand-name drugs are supplied to a segment of the continuum, that is, $\theta_C - \underline{\theta}$ where $\underline{\theta} < \theta_C < 1$ with θ_C endogenously chosen.

With the subscripts or superscripts A , B , and C referring to the three cases and capital letters denoting aggregate levels, the market demand functions for brand-name drugs in each of the above three cases are

$$X_b^A = \int_{\underline{\theta}}^1 \left(\frac{\alpha}{\beta} - \frac{\theta p_b}{\beta} \right) d\theta = (1 - \underline{\theta}) \frac{\alpha}{\beta} - \left(\frac{1 - \underline{\theta}^2}{2} \right) \frac{p_b^A}{\beta}, \quad (7)$$

$$X_b^B = \int_{\underline{\theta}}^{\theta_B} \left(\frac{\alpha}{\beta} - \frac{\theta p_b^B}{\beta} \right) d\theta = (\theta_B - \underline{\theta}) \frac{\alpha}{\beta} - \left(\frac{\theta_B^2 - \underline{\theta}^2}{2} \right) \frac{p_b^B}{\beta}, \quad (8)$$

and

$$\begin{aligned} X_b^C &= \int_{\underline{\theta}}^{\theta_C} \left\{ \frac{\alpha + (\beta + \gamma) q}{\beta + \gamma} - \left(\frac{1}{\beta^2 - \gamma^2} \right) [\beta \theta p_b^C - \gamma (1 - t) \theta p_g^C] \right\} d\theta \\ &= (\theta_C - \underline{\theta}) \left[\frac{\alpha + (\beta + \gamma) q}{\beta + \gamma} \right] - \frac{1}{2} \left(\frac{\theta_C^2 - \underline{\theta}^2}{\beta^2 - \gamma^2} \right) [\beta p_b^C - \gamma (1 - t) p_g^C]. \end{aligned} \quad (9)$$

In CASE C, the market demand function for generic drugs is

$$\begin{aligned} X_g^C &= \int_{\underline{\theta}}^1 \left\{ \frac{\alpha - (\beta + \gamma) q}{\beta + \gamma} - \left(\frac{1}{\beta^2 - \gamma^2} \right) [\beta (1 - t) \theta p_g^C - \gamma \theta p_b^C] \right\} d\theta \\ &= (1 - \underline{\theta}) \left[\frac{\alpha - (\beta + \gamma) q}{\beta + \gamma} \right] - \frac{1}{2} \left(\frac{1 - \underline{\theta}^2}{\beta^2 - \gamma^2} \right) [\beta (1 - t) p_g^C - \gamma p_b^C]. \end{aligned} \quad (10)$$

⁸When generic drugs are not available (CASES A and B), type- θ consumer's utility function, upon incorporation of his/her budget constraint, is given by

$$U(\cdot) = m - \theta p_b x_b + \alpha x_b - \frac{1}{2} \beta x_b^2$$

and his/her demand for brand-name drugs is

$$x_b = \frac{\alpha - \theta p_b}{\beta}.$$

When segmentation occurs (CASES B and C), the price of brand-name drugs is such that brand-name drugs are not purchased by consumers with $\theta \in [\theta_B, 1]$ in CASE B and by consumers with $\theta \in [\theta_C, 1]$ in CASE C; that is, in each of the two segmentation cases, individuals with low insurance coverage (high θ) are “priced out of the market” for brand-name drugs.

To be able to more accurately understand the relevance of incorporating differential health insurance coverages into the analysis of pricing responses to generic entry, the equilibrium is also derived in the absence of insurance considerations when only brand-name drugs are available (CASE D) and when both brand-name and generic drugs are available (CASE E); in both cases, the entire market consisting of $1 - \underline{\theta}$ consumers is served. Hence, for CASE D, the market demand function for brand-name drugs is

$$X_b^D = (1 - \underline{\theta}) \left(\frac{\alpha - p_b^D}{\beta} \right) \quad (11)$$

and, for CASE E, the market demand functions for brand-name and generic drugs are

$$X_b^E = (1 - \underline{\theta}) \left[\frac{\alpha + (\beta + \gamma)q}{\beta + \gamma} - \left(\frac{1}{\beta^2 - \gamma^2} \right) (\beta p_b^E - \gamma p_g^E) \right] \quad (12)$$

and

$$X_g^E = (1 - \underline{\theta}) \left[\frac{\alpha - (\beta + \gamma)q}{\beta + \gamma} - \left(\frac{1}{\beta^2 - \gamma^2} \right) (\beta p_g^E - \gamma p_b^E) \right]. \quad (13)$$

2.1 The Equilibrium with Homogenous Consumers

When consumers enjoy the same health insurance coverage, they can be assumed to have no coverage without any loss of generality (that is, $\theta = 1$). With homogenous consumers, market segmentation is impossible and drugs' producers supply to the entire market of $1 - \underline{\theta}$ consumers. Under the assumption that the marginal cost of production, denoted by c , is constant with $c < \alpha$, two cases are considered: CASE D in which only brand-name drugs are produced; CASE E in which both brand-name and generic drugs are produced.

2.1.1 CASE D: No Generic Drugs, Full Market Coverage, and No Insurance Coverage

In the absence of insurance coverage, the profit function, upon substitution for p_b^D from (11), can be written as

$$\pi_b^D = \left[\alpha - \left(\frac{1}{1-\underline{\theta}} \right) X_b^D - c \right] X_b^D, \quad (14)$$

the maximization of which with respect to X_b^D gives

$$p_b^D = \frac{\alpha + c}{2}. \quad (15)$$

2.1.2 CASE E: Generic Drugs, Full Market Coverage, and No Insurance Coverage

When generic drugs become available, the supply side of the pharmaceutical market is characterized by n quantity-competing firms which produce generic drugs and a monopolist which produces brand-name drugs. Each firm faces the same constant marginal cost (c) independently of the type of drugs it produces. The model is then a two-stage game: in the first stage, the monopolist chooses how many types of consumers to supply to; in the second stage, the n producers of generic drugs engage in Cournot competition taking the price of brand-name drugs as given. To derive the equilibrium, the second stage of the model is considered first to obtain the price of generic drugs as a function of the price of brand-name drugs; upon substitution of this price into the profit function of the monopolist, the price of the brand-name drugs is determined.

The Second Stage of the Game: Generic Firms' Cournot Competition

Using (13) to obtain p_g^E and that $X_g^E = nx_{gi}^E$, the profit function of firm i (for $i = 1, \dots, n$) is given by

$$\pi_{gi}^E = \left\{ \frac{[\alpha - (\beta + \gamma)q](\beta - \gamma)}{\beta} + \frac{\gamma}{\beta} p_b^E - \frac{n(\beta^2 - \gamma^2)}{(1-\underline{\theta})\beta} x_{gi}^E - c \right\} x_{gi}^E, \quad (16)$$

which is maximized when

$$p_g^E = \frac{[\alpha - (\beta + \gamma)q](\beta - \gamma)}{2\beta} + \frac{c}{2} + \frac{\gamma}{2\beta} p_b^E. \quad (17)$$

The First Stage of the Game: Brand-Name Firm's Price Setting Upon substitution for p_g^E from (17) into (12) and rearrangement to obtain p_b^E , the monopolist's profit function can be written as

$$\begin{aligned} \pi_b^E = & \left\{ \frac{[(2\beta + \gamma)\alpha + (\beta + \gamma)(2\beta - \gamma)q](\beta - \gamma) + \gamma c}{2\beta^2 - \gamma^2} \right\} X_b^E - \\ & - \left[\frac{2(\beta^2 - \gamma^2)\beta}{(1 - \theta)(2\beta^2 - \gamma^2)} X_b^E + c \right] X_b^E. \end{aligned} \quad (18)$$

Maximization of the above profit function yields

$$p_b^E = \frac{[c + (\beta - \gamma)q](\beta + \gamma)(2\beta - \gamma) + \alpha(\beta - \gamma)(2\beta + \gamma)}{2(2\beta^2 - \gamma^2)}. \quad (19)$$

2.1.3 Comparative Analysis of Prices

In the absence of insurance coverage, a comparison of the prices of brand-name drugs before and after generic entry (that is, p_b^D and p_b^E) leads to the following proposition:

Proposition 1 *In the absence of health insurance coverage considerations, $p_b^E > p_b^D$ iff $q > \frac{\beta\gamma(\alpha - c)}{(2\beta - \gamma)(\beta^2 - \gamma^2)}$.*

When heterogeneity in health insurance coverage is ignored, the price of brand-name drugs increases following the introduction of generic drugs provided that the perceived quality differential between the two types of drugs is sufficiently large. More precisely, ceteris paribus, the lower the degree of product differentiation between brand-name and generic drugs (γ/β) and/or the higher the perceived quality differential between brand-name and generic drugs (q), the more likely it is for the price of brand-name drugs to increase in response to the entry of generic drugs. The above result follows directly from a comparison of p_b^D in (15) and p_b^E in (19).

Although consumers are assumed to be homogenous so that no segmentation can occur, the inclusion of q into the utility function allows for a re-consideration of the brand loyalty argument proposed by Caves et al. (1991), Grabowski and Vernon (1992), and Frank and Salkever (1997) to explain the generic competition paradox. In essence, brand loyalty arises because consumers perceive brand-name products to be of better quality; by capturing consumers' perception of the quality superiority of the brand-name drugs, q does reflect, and is

thus a proxy for, brand loyalty. The above proposition therefore supports the brand loyalty story to justify the price increase for brand-name drugs when competition in the pharmaceutical market intensifies as a result of generic entry. If brand loyalty as captured by q is significant relative to some measure of market profitability as captured by c/α and/or relative to the degree of product differentiation between the two types of drugs (that is, γ/β), the generic competition paradox is likely to occur. Interestingly, the presence of brand loyalty is a necessary but not a sufficient condition for the paradox; in other words, it is possible for generic entry to entail a non-positive change in the price of brand-name drugs even when consumers exhibit a high degree of brand loyalty provided that the two types of drugs are close substitutes (high γ/β) and/or the industry is more profitable (low c/α).⁹

2.2 The Equilibrium with Heterogenous Consumers

When health insurance coverage is included into the model so that consumers can be allowed to differ, the price effects of market segmentation based on insurance coverage can be assessed in instances in which brand loyalty is insufficient to generate the GCP. For completeness, in the absence of generic drugs, the profit-maximizing problem the monopolist producing brand-name drugs faces is considered both when the entire market of $1 - \underline{\theta}$ consumers is covered (CASE A), in which case the choice variable is X_b , and when market segmentation based on insurance coverage occurs (CASE B), in which case the choice variable is θ_B (that is, the lowest acceptable coverage for the purchase of brand-name drugs). The market-segmentation case is examined when generic drugs are not available to separate between the effect (on the price of brand-name drugs) of market segmentation and the effect of generic entry through market segmentation.

In the presence of generic drugs, the producer of brand-name drugs moves first and decides on how to segment the market or how many types of consumers to supply to (that is, it chooses θ_C , with $\underline{\theta} < \theta_C < 1$, so that only $\theta_C - \underline{\theta}$ consumers buy brand-name drugs or consumers with $\theta < \theta_C$). The n quantity-competing producers of generic drugs move next

⁹The parameter α is a determinant of the maximum willingness to pay for drugs. The higher the α , the more consumers are willing to pay for drugs. Hence, the difference between α and c is a measure of the surplus firms can potentially extract from consumers and the ratio of c to α is thus a measure of market profitability, with a lower ratio corresponding to a larger gap between α and c .

and, given the price of brand-name drugs, determine how much to produce taking the entire continuum of consumers into account, of which $\theta_C - \underline{\theta}$ buy both types of drugs and $1 - \theta_C$ buy only generic drugs.

2.2.1 CASE A: No Generic Drugs and Full Market Coverage

Substituting for p_b^A from (7) into the profit function gives

$$\pi_b^A = \left[\frac{2\alpha}{(1 + \underline{\theta})} - \left(\frac{2\beta}{1 - \underline{\theta}^2} \right) X_b^A - c \right] X_b^A, \quad (20)$$

the maximization of which with respect to X_b^A yields

$$p_b^A = \frac{\alpha}{(1 + \underline{\theta})} + \frac{c}{2} > p_b^D. \quad (21)$$

The higher the marginal willingness to pay for drugs (α) and/or the higher the marginal production cost (c) and/or the larger the market ($1 - \underline{\theta}$), the higher the price of brand-name drugs is. The relation between p_b^D and p_b^A is straightforward as $p_b^D - p_b^A = -\frac{\alpha}{2} \left(\frac{1 - \underline{\theta}}{1 + \underline{\theta}} \right) < 0$; hence, the presence of varying degrees of coverage results in an increase in the price of brand-name drugs when generic drugs are not available in the market and the entire market is serviced. Insurance coverage effectively reduces the out-of-pocket cost of healthcare, thus increasing the willingness to pay for healthcare and enabling the monopolist to sell more at a higher price.

2.2.2 CASE B: No Generic Drugs and Partial Market Coverage

In this case, the producer of brand-name drugs supplies to only a segment of the market; specifically, it sets p_b^B such that type- θ_B consumer buys nothing. From footnote 8, the type- θ_B consumer's demand for brand-name drugs is given by $x_b^B = \frac{\alpha - \theta_B p_b^B}{\beta}$ so that $p_b^B = \frac{\alpha}{\theta_B}$. Upon substitution for p_b^B into (8), the profit function is

$$\pi_b^B = \left(\frac{\alpha}{\theta_B} - c \right) \left[(\theta_B - \underline{\theta}) - \frac{1}{2} \left(\frac{\theta_B^2 - \underline{\theta}^2}{\theta_B} \right) \right] \frac{\alpha}{\beta}. \quad (22)$$

Maximization of (22) with respect to θ_B yields

$$\alpha \left(\frac{\theta_B + \underline{\theta}}{\theta_B} \right) - \frac{c}{2} (\theta_B + \underline{\theta}) = \alpha, \quad (23)$$

which gives

$$\theta_B = -\frac{\theta}{2} + \sqrt{\frac{\theta^2}{4} + \frac{2\alpha\theta}{c}} \quad (24)$$

and

$$p_b^B = \frac{\alpha}{-\frac{\theta}{2} + \sqrt{\frac{\theta^2}{4} + \frac{2\alpha\theta}{c}}} > p_b^A. \quad (25)$$

For an interior solution, θ_B must range between $\underline{\theta}$ and 1 and this requires that $\frac{\alpha}{c} < \frac{1+\underline{\theta}}{2\underline{\theta}}$.¹⁰ That $p_b^B > p_b^A$ for $\theta_B < 1$ follows directly from (23) which gives that $\frac{\alpha}{\theta_B} - \frac{c}{2} = \frac{\alpha}{(\theta_B + \underline{\theta})}$; hence, $p_b^B - p_b^A = \frac{\alpha}{\theta_B} - \frac{\alpha}{(1+\underline{\theta})} - \frac{c}{2} = \frac{\alpha}{(\theta_B + \underline{\theta})} - \frac{\alpha}{(1+\underline{\theta})} > 0$. The only parameters affecting θ_B are $\underline{\theta}$ and the ratio of c to α , which is relabeled as f with $0 < f < 1$. In particular, comparative statics lead to the following results:

Proposition 2 *The higher $\underline{\theta}$ and/or the lower c/α , the larger θ_B is and the smaller p_b^B is.*

A higher $\underline{\theta}$ corresponds to a worsening in the best available insurance coverage, which is equivalent to a decrease in the size of the market; a lower f , where $f = c/\alpha$, corresponds to an increase in market profitability. Hence, a decrease in the market size ($\underline{\theta} \uparrow$) induces the monopolist to provide brand-name drugs to consumers with lower insurance coverage ($\theta_B \uparrow$), thus charging a lower price for brand-name drugs ($p_b^B \downarrow$). The segment of the market the monopolist supplies to may increase or decrease depending on the value of the product between $\underline{\theta}$ and f ; specifically, the segment increases at low values of $\underline{\theta}f$ (and vice versa). Furthermore, a decrease in market profitability ($f \uparrow$) induces the monopolist to supply to fewer types of consumers, or to supply to consumers with better insurance coverage ($\theta_B \downarrow$), and to charge a higher price for brand-name drugs ($p_b^B \uparrow$).

Partial differentiation of (24) with respect to $\underline{\theta}$ and f , where $f = \frac{c}{\alpha}$, gives

$$\frac{\partial \theta_B}{\partial \underline{\theta}} = \frac{\underline{\theta}f + 4 - \sqrt{\underline{\theta}f(\underline{\theta}f + 8)}}{2\sqrt{\underline{\theta}f(\underline{\theta}f + 8)}} > 0 \quad (26)$$

and

$$\frac{\partial \theta_B}{\partial f} = -\frac{\underline{\theta}}{f^2 \sqrt{\frac{\theta^2}{4} + \frac{2\theta}{f}}} < 0. \quad (27)$$

¹⁰For $\theta_B > \underline{\theta}$, the condition that $\frac{\alpha}{c} > \underline{\theta}$ is satisfied as $\alpha > c$ and $\underline{\theta} < 1$.

For $\underline{\theta}f < 3\sqrt{2} - 4$, the increase in θ_B is larger than the increase in $\underline{\theta}$ so that, overall, the monopolist ends up servicing more types of consumers. From footnote 8, the maximum consumer surplus that could be extracted from type- θ consumer at a constant marginal cost of c is given by $\alpha^2 (1 - \theta f)^2 / 2\theta\beta$, which is a decreasing function of θ , so that $(1 - \theta f)^2$ is equal to the ratio of consumer surplus at c to consumer surplus at zero marginal cost under marginal cost pricing. In other words, $(1 - \theta f)^2$ provides a measure of the loss in the realizable benefit of servicing type- θ consumer when production entails a positive marginal cost equal to c . Specifically, for $(1 - \theta f)^2 > (5 - 3\sqrt{2})^2 \approx 0.58$ (or the consumer surplus that can be extracted from type- θ consumer at c is at least 58 percent of the consumer surplus that can be extracted at zero marginal cost), $\frac{\partial \theta_B}{\partial \underline{\theta}} > 1$ and more types of consumers buy brand-name drugs; that is, the monopolist chooses to supply to more types of consumers in response to a decrease in market size or a worsening in the best insurance coverage. From (27), as f decreases, the loss in the surplus that can be extracted from the consumer with no coverage ($\theta = 1$) when production is costly decreases, thus inducing the monopolist to service more types of consumers.¹¹

2.2.3 CASE C: Generic Drugs and Partial Market Coverage

When generic drugs become available, the model involves a two-stage game with one producer of brand-name drugs and n producers of generic drugs: in the first stage, the segment of the market which buys both types of drugs and the price of brand-name drugs are determined; in the second stage, the price of generic drugs is determined.

The Second Stage of the Game: Generic Firms' Cournot Competition The profit of firm i (for $i = 1, \dots, n$) is given by

$$\pi_{gi}^C = (p_g^C - c) x_{gi}^C, \quad (28)$$

where

$$p_g^C = \frac{2[\alpha - (\beta + \gamma)q](\beta + \gamma)}{(1 - t)(1 + \underline{\theta})\beta} + \frac{\gamma}{(1 - t)\beta} p_b^C - \frac{2(\beta^2 - \gamma^2)}{(1 - t)(1 - \underline{\theta}^2)\beta} X_g^C \quad (29)$$

¹¹For $\theta = 1$, the maximum extractable consumer surplus at c as a ratio of the maximum extractable consumer surplus at zero marginal cost is in fact given by $(1 - f)^2$ so that $1 - (1 - f)^2 = f(2 - f)$ gives the percentage of the maximum willingness to pay of the consumer with no coverage at zero marginal cost that is lost when marginal cost is a positive constant equal to c .

from (10). As the n firms are identical, they produce the same quantity in equilibrium so that $x_{gi}^C = \frac{X_g^C}{n}$ and the market-clearing price of generic drugs as a function of the price of brand-name drugs is given by

$$p_g^C = \frac{[\alpha - (\beta + \gamma)q](\beta - \gamma)}{(1-t)(1+\underline{\theta})\beta} + \frac{c}{2} + \frac{\gamma}{2(1-t)\beta} p_b^C. \quad (30)$$

The First Stage of the Game: Brand-Name Firm's Price Setting With θ_C denoting the insurance parameter segmenting the market between consumers buying both types of drugs, $\theta_C - \underline{\theta}$, and consumers buying only generic drugs, $1 - \theta_C$, the monopolist's problem is to maximize

$$\pi_b^C = (p_b^C - c) X_b^C \quad (31)$$

with respect to θ_C taking into account how p_g^C responds to p_b^C as captured by (30), where

$$p_b^C = \frac{[\alpha + (\beta + \gamma)q](\beta - \gamma)}{\beta\theta_C} + \frac{\gamma(1-t)}{\beta} p_g^C \quad (32)$$

from (5) to ensure that type- θ_C consumer does not buy brand-name drugs. The condition for an interior solution to the above maximization problem can then be expressed as

$$A\theta_C^2 + A\underline{\theta}\theta_C + B = 0, \quad (33)$$

so that

$$\theta_C = -\frac{\underline{\theta}}{2} + \sqrt{\frac{\underline{\theta}^2}{4} - \frac{B}{A}}, \quad (34)$$

where

$$A = \frac{2[\alpha - (\beta + \gamma)q](\beta - \gamma)\gamma}{(1+\underline{\theta})} + c[\beta\gamma(1-t) - 2\beta^2 + \gamma^2] < 0 \quad (35)$$

and

$$B = 4\beta[\alpha + (\beta + \gamma)q](\beta - \gamma)\underline{\theta} > 0. \quad (36)$$

The equilibrium price is thus given by

$$\begin{aligned} p_b^C &= \left(\frac{2\beta}{2\beta^2 - \gamma^2} \right) \left\{ \frac{[\alpha - (\beta + \gamma)q](\beta - \gamma)\gamma}{(1+\underline{\theta})\beta} + \frac{c\gamma(1-t)}{2} \right\} + \\ &+ \left(\frac{2\beta}{2\beta^2 - \gamma^2} \right) \left\{ \frac{[\alpha + (\beta + \gamma)q](\beta - \gamma)}{\theta_C} \right\}, \end{aligned} \quad (37)$$

with θ_C as in (34).

2.2.4 Comparative Analysis of Market Segmentation and Price Setting

In order to isolate the contribution of incorporating heterogenous health insurance coverages to explaining the generic competition paradox, the brand loyalty argument as above defined is assumed away. In particular, the brand loyalty parameter (q) is set equal to $\frac{\beta\gamma(\alpha-c)}{(2\beta-\gamma)(\beta^2-\gamma^2)}$ such that the price of brand-name drugs is not affected by generic entry in the absence of heterogenous health insurance coverage (that is, $p_b^D = p_b^E$).

With s denoting the value of γ/β , that is, the degree of product differentiation between brand-name and generic drugs, which ranges between 0 (products are independent) and 1 (products are perfect substitutes),¹² and f denoting the value of c/α , for $0 < \frac{2\theta}{(1+\theta)} < f < 1$, the above condition becomes $q = \frac{s\alpha(1-f)}{\beta(2-s)(1-s^2)}$ and the price of brand-name drugs in CASE C (that is, when generic drugs are available to the entire market but brand-name drugs are only supplied to consumers with an insurance coverage parameter ranging between θ_C and $\underline{\theta}$) simplifies to

$$p_b^C = \alpha \left(\frac{2}{2-s^2} \right) \left\{ \frac{\left[1 - s - \frac{s(1-f)}{(2-s)} \right] s}{(1+\underline{\theta})} + \frac{fs(1-t)}{2} + \frac{\left[1 - s + \frac{s(1-f)}{(2-s)} \right]}{\theta_C} \right\}, \quad (38)$$

where

$$\theta_C = -\frac{\theta}{2} + \sqrt{\frac{\theta^2}{4} - \frac{N}{D}}, \quad (39)$$

with

$$N = 4\underline{\theta}(1+\underline{\theta})[(2-s)(1-s) + s(1-f)] > 0 \quad (40)$$

and

$$D = f(1+\underline{\theta})(2-s)[(1-t)s - (2-s^2)] + 2s[(2-s)(1-s) - s(1-f)], \quad (41)$$

which must be negative for an interior solution of θ_C .¹³ The extent of segmentation as captured by θ_C depends not only upon $\underline{\theta}$ and f , as in CASE B, but also on t and s . The specifics of this dependence are given in the following proposition:

¹²More precisely, $0 \leq s < 1$ as, by assumption, $\beta > \gamma$.

¹³For an interior solution of θ_C , $f_{\min} < f < f_{\max}$, where $f_{\min} = \max\left(\frac{2\theta}{1+\underline{\theta}}, f_L\right)$ and $f_{\max} = \min(f_H, 1)$. See appendix for details on f_L and f_H .

Proposition 3 *A decrease in t increases θ_C as does an increase in $\underline{\theta}$. An increase in f is likely to reduce θ_C while an increase in s reduces θ_C at high values of s and increases it at low values of s .*

As the additional deductible on the purchase on generic drugs decreases ($t \downarrow$) and/or the market size decreases ($\underline{\theta} \uparrow$), the monopolist opts to increase θ_C , thus supplying to more types of consumers. Furthermore, as the cost of producing drugs increases ($c \uparrow$) or the willingness to pay for drugs decreases ($\alpha \downarrow$), causing f , which is defined as $\frac{c}{\alpha}$, to increase, the producer of brand-name drugs chooses to lower θ_C , thus supplying to fewer types of consumers, independently of the degree of product differentiation between brand-name and generic drugs (i.e., s), provided that the additional insurance coverage for generic drugs (i.e., t) is not sufficiently small. Finally, as the degree of product differentiation between the two types of drugs increases ($s \uparrow$), the monopolist supplies to more (fewer) consumers when the degree of product differentiation is lower (larger) and the industry is profitable (i.e., f is low); for a given t and $\underline{\theta}$, the relevant range of high (low) values of s over which the monopolist supplies to fewer (more) consumers in response to an increase in s gets narrower (wider) as f increases up to some critical level and vice versa as f increases past this critical level.

Partial differentiation of θ_C from (39) with respect to t and $\underline{\theta}$ yields

$$\frac{\partial \theta_C}{\partial t} = -\frac{f(1+\underline{\theta})(2-s)sN}{(2\theta_C+\underline{\theta})D^2} < 0 \quad (42)$$

and

$$\frac{\partial \theta_C}{\partial \underline{\theta}} = \frac{(1+\underline{\theta})\theta_C^2 - 2\underline{\theta}s[(2-s)(1-s) - s(1-f)]\frac{N}{D^2}}{\underline{\theta}(1+\underline{\theta})(2\theta_C+\underline{\theta})} > 0. \quad (43)$$

Partial differentiation of θ_C from (39) with respect to f gives

$$\frac{\partial \theta_C}{\partial f} = \frac{4\underline{\theta}(1+\underline{\theta})(2-s)\{(1+\underline{\theta})(2-2s+s^2)[(1-t)s - (2-s^2)] + 4s^2(1-s)\}}{(2\theta_C+\underline{\theta})D^2}, \quad (44)$$

which is negative for $0 < s \leq 0.85$ independently of the values of $\underline{\theta}$ and t and for $0.86 \leq s < 1$ provided that $\underline{\theta}$ and t are such that $ts(1+\underline{\theta}) - \underline{\theta}(s^2+s-2) > \frac{(1-s)(-s^3+4s^2+2s-4)}{s^2-2s+2}$ (this

condition is always satisfied for $t \geq 0.04$).^{14,15} Finally, partial differentiation of θ_C from (39) with respect to s gives

$$\frac{\partial \theta_C}{\partial s} = -\frac{\Phi}{(2\theta_C + \underline{\theta}) D^2}, \quad (45)$$

where $\Phi = D \frac{\partial N}{\partial s} - N \frac{\partial D}{\partial s}$ which is equal to zero for

$$t = \frac{s(2 - f\underline{\theta})(s^3 - 4s^2 + 12s - 12) - 4(2s - 5sf - 2) - s^2 f(s^2 + 8f - 2f^2)}{f(1 + \underline{\theta})[4(s - 1) - fs^2]} + \frac{f(2s^3 - s^2 - 4)}{[4(s - 1) - fs^2]}. \quad (46)$$

For given t and $\underline{\theta}$, (46) defines the isovalue curve for $\Phi = 0$ in (f, s) space such that θ_C decreases in response to an increase in s above the curve but increases below the curve. The isovalue curve for $t = \underline{\theta} = 0.10$ is illustrated in Figure 1. Ceteris paribus, an increase in s results in an increase in Φ as

$$\left. \frac{\partial t}{\partial s} \right|_{\Phi=0} = -\frac{2[4f + fs^3 - 6 - 6(1 - s)^2][(1 - s)^2 + (1 - sf)][f(1 + \underline{\theta}) - 2]}{f(1 + \underline{\theta})(4 - 4s + fs^2)^2} < 0, \quad (47)$$

that is, an increase in s has to be accompanied by a decrease in t in order for Φ to remain at zero, and

$$\frac{\partial \Phi}{\partial t} = 4\underline{\theta}(1 + \underline{\theta})^2 f(fs^2 - 4s + 4) > 0, \quad (48)$$

so that $\Phi > 0$ ($\Phi < 0$) above (below) a given $t = k$ isovalue curve, for $0 \leq k \leq 1$. At values of f that yield interior solutions for θ_C , that is, for $f > f_L$ as defined in footnote 12, it then follows that $\frac{\partial \theta_C}{\partial s} < 0$ above the isovalue curve (dark grey area) and $\frac{\partial \theta_C}{\partial s} > 0$ below the isovalue curve (light grey area); in other words, the market share of brand-name drugs decreases (increases) as the two types of drugs become less differentiated at high (low) values of s .

From (47), the isovalue curve for $t = k$ shifts down as k increases; furthermore, from

$$\left. \frac{\partial t}{\partial \underline{\theta}} \right|_{\Phi=0} = -\frac{2\{s^4 - 2(2 + f)s^3 + [12 + f(2 - f)]s^2 - 4(4 - f)s + 4\}}{f(1 + \underline{\theta})^2(4 - 4s + fs^2)}, \quad (49)$$

¹⁴In the absence of a preferential treatment of generic drugs by insurance companies through a higher coverage (that is, when $t = 0$), $\underline{\theta} > \frac{(-s^3 + 4s^2 + 2s - 4)}{(s + 2)(s^2 - 2s + 2)}$ for an increase in f to result into a decrease in θ_C ; put differently, at low values of $\underline{\theta}$ and high values of s ($0.02 \leq \underline{\theta} \leq 0.32$ for $0.86 \leq s \leq 0.99$), the monopolist opts to supply to fewer types of consumers in response to an increase in production costs and/or a decrease in the willingness to pay for drugs.

¹⁵See appendix for more details on $\frac{\partial \theta_C}{\partial \underline{\theta}}$ and $\frac{\partial \theta_C}{\partial f}$.

which is negative for $f < \hat{f}$, the isovalue curve for $t = k$ shifts up as $\underline{\theta}$ increases for $f < 1 - k$ and down for $f > 1 - k$. For a given $\underline{\theta}$ (e.g., $\underline{\theta} = 0.1$ in Figure 1), the $t = k$ isovalue curve slopes upward for $f < \tilde{f}$, where \tilde{f} is such that $\left. \frac{\partial t}{\partial f} \right|_{\Phi=0} = 0$, and downward for $f > \tilde{f}$; hence, the range of low values of s over which $\frac{\partial \theta_C}{\partial s} > 0$ gets wider as f increases up to $f = \tilde{f}$ and narrower for $f > \tilde{f}$.¹⁶ The effects of changes in $\underline{\theta}$ and t on how θ_C responds to a change in s are illustrated in Figure 1a and Figure 1b. In Figure 1a, the $t = 0.10$ isovalue curve shifts up when $\underline{\theta}$ increases from 0.10 to 0.30 for $0 < f < 0.90$ and it shifts down for $0.90 < f < 1$; the increase in $\underline{\theta}$ also affects f_L making it less likely for an interior solution to result at combinations of low values of f and s . Overall, the region in which $\Phi > 0$, so that $\frac{\partial \theta_C}{\partial s} < 0$, decreases by A_1 and the light grey area for $f < 0.90$ and it increases by the dark grey area for $f > 0.90$; the region in which $\Phi < 0$, so that $\frac{\partial \theta_C}{\partial s} > 0$, decreases by A_2 but increases by the light grey area for $f < 0.90$ and it decreases by the dark grey area for $f > 0.90$ (in A_1 and A_2 , everyone buys brand-name drugs, that is, $\theta_C = 1$). In Figure 1b, both the f_L and the $t = k$ isovalue curves shift down when k increases from 0.10 to 0.30; the region in which $\Phi < 0$ thus increases by the dark grey area and A_1 while the region in which $\Phi > 0$ decreases by the dark grey area and increases by A_2 (in A_1 and A_2 , everyone buys brand-name drugs when $t = 0.10$ but not when $t = 0.30$).¹⁷

In terms of market segmentation, the effect of heterogenous insurance coverage is ambiguous; however, the following result obtains:

Proposition 4 *For given $\underline{\theta}$ and t , there exists an s^* corresponding to each value of f such that $\theta_B > \theta_C$ for $s > s^*$ and $\theta_C > \theta_B$ for $s < s^*$. As t increases, s^* decreases; as $\underline{\theta}$ increases, s^* increases for $f < f^*$ and decreases for $f > f^*$, where $f^* = 1 - t$.*

When segmentation based on insurance coverage occurs in the absence of generic drugs, the market share of brand-name drugs may increase in response to generic entry. Specifically, ceteris paribus, brand-name drugs are made accessible to more consumers (and thus to consumers with lower coverage) following the introduction of generic drugs when the two

¹⁶See appendix for details on \hat{f} , \tilde{f} , and the $t = k$ isovalue curve.

¹⁷An increase in $\underline{\theta}$ increases f_L while an increase in t decreases f_L . In terms of the isovalue curve in (f, s) space such that $f = f_L$ for given $\underline{\theta}$ and t , an increase in $\underline{\theta}$ shifts the curve out while an increase in t shifts the curve in. See appendix.

types of drugs are more differentiated; as the pharmaceutical market becomes less profitable ($c \uparrow$ or $\alpha \downarrow$ or both so that $f \uparrow$), generic drugs can be increasingly closer substitutes for brand-name drugs, at least up to some critical value of f , and still trigger an increase in the market share of brand-name drugs upon entry. However, the more favourable the co-payment system is on the purchase of generic drugs relative to brand-name drugs ($t \uparrow$), the more likely it is for some consumers to switch completely to generic drugs; on the other hand, the smaller the market is ($\underline{\theta} \uparrow$), the more (less) likely it is for some consumers to go from buying nothing before generic entry to buying brand-name drugs after generic entry for $f < 1 - t$ ($f > 1 - t$).

With M denoting the difference between $-\frac{N}{D}$ as above defined (this ratio is equal to $-\frac{B}{A}$ from (35) and (36) written in terms of f and s under the restriction that q is such that the price of brand-name drugs does not change in response to generic entry when insurance coverage considerations are omitted from the analysis) and $\frac{2\theta}{f}$, so that $\theta_C > \theta_B$ if $M > 0$ and vice versa, it obtains that $M \leq 0$ for $s \geq s^*$, where

$$s^* = \frac{R + f[(1 + \underline{\theta})(3 + t) + 2] - 8}{2[f(1 + \underline{\theta}) - 2]}, \quad (50)$$

with $R > 0$, which is a concave function of f as illustrated in Figure 2.¹⁸ When $f = f^* = 1 - t$, $s^* = \frac{3+t-\sqrt{t^2+6t+1}}{2}$ and is thus independent of $\underline{\theta}$ (e.g., in Figure 2a, $s^* = 0.9156$ at $f^* = 0.90$ for any of the values of $\underline{\theta}$ considered). For given $\underline{\theta}$ and t , there is an s^* curve below which $\theta_C > \theta_B$ and above which $\theta_B > \theta_C$; the s^* curve reaches its maximum at $s = 2 - \sqrt{2 - f^2(1 + \underline{\theta})}$ which is a positive function of f as drawn in Figure 2. The effects of changes in $\underline{\theta}$ and t on s^* , which are depicted in Figures 2a and 2b, can then be expressed as

$$\frac{\partial s^*}{\partial \underline{\theta}} = -\frac{f(f - 1 + t)[ft(1 + \underline{\theta}) + f(1 - \underline{\theta}) + R]}{[f(1 + \underline{\theta}) - 2]^2 R}, \quad (51)$$

which is positive for $f < f^* = 1 - t$ and negative otherwise, and

$$\frac{\partial s^*}{\partial t} = \frac{f(1 + \underline{\theta})[ft(1 + \underline{\theta}) + f(1 - \underline{\theta}) + R]}{2[f(1 + \underline{\theta}) - 2]R} > 0. \quad (52)$$

Graphically, an increase in $\underline{\theta}$ rotates the s^* curve clockwise around f^* so that the s^* curve moves up for $f < f^*$ and down for $f > f^*$ (Figure 2a); an increase in t shifts the entire

¹⁸See appendix.

s^* curve down (Figure 2b). With the lower-bound of f (f_{\min}) also dependent on $\underline{\theta}$ and t ,¹⁹ changes in $\underline{\theta}$ and t create or remove areas in which only corner solutions are feasible. In terms of Figure 2a, for example, an increase $\underline{\theta}$ from 0.1 to 0.3 results in the following changes: in A_1 , the relationship $\theta_B > \theta_C$ still holds but $\theta_B = 1$; in A_2 , the relationship $\theta_B > \theta_C$ no longer holds and $\theta_B = \theta_C = 1$; in A_3 , the relationship $\theta_C > \theta_B$ no longer holds and $\theta_B = \theta_C = 1$; in A_4 , the relationship $\theta_C > \theta_B$ still holds but $\theta_C = 1$; in A_5 and A_6 , the relationship shifts from $\theta_B > \theta_C$ to $\theta_C > \theta_B$ but $\theta_C = 1$ in A_5 ; in A_7 , the relationship shifts from the $\theta_C > \theta_B$ to $\theta_B > \theta_C$. In terms of Figure 2b, an increase in t from 0.1 to 0.2 triggers the following changes: in A_1 , the relationship shifts from the $\theta_C > \theta_B$ to $\theta_B > \theta_C$; in A_2 , the relationship $\theta_C > \theta_B$ still holds but θ_C is no longer equal to 1.

In general, the market share of brand-name drugs increases in response to generic entry when s is relatively low and decreases when s is relatively high. In Figures 2c and 2d, θ_C and θ_B are drawn as functions of f for $t = \underline{\theta} = 0.10$ when the two types of drugs are considered weak substitutes (e.g., $s = 0.5$ in Figure 3a) and when they are considered strong substitutes (e.g., $s = 0.9$ in Figure 3b). Consistent with the results summarized in Figure 2, $\theta_C > \theta_B$, which implies that the market share of brand-name drugs increases in response to the entry of generic drugs, at any value of f when s is relatively low (Figure 3a), and $\theta_C < \theta_B$ at both high and low values of f when s is relatively high (Figure 3b). In both Figures, the light grey area captures the extent to which the market share of brand-name drugs increases as a function of f with the lower-bound value of f given by 0.41 (that is, f_L as defined in footnote 12) in Figure 3a and by 0.18 (that is, $\frac{2\underline{\theta}}{(1+\underline{\theta})}$) in Figure 3b; the dark grey area in Figure 3b captures the extent to which the market share of brand-name drugs declines following generic entry.

Upon comparison of p_b^C with p_b^A (that is, the prices of brand-name drugs after and before the introduction of generic drugs under the assumption that the entire market is served when only brand-name drugs are available) and of p_b^C with p_b^B (that is, the prices of brand-name drugs after and before the introduction of generic drugs under the assumption that

¹⁹The lower-bound of f is equal to the $\max\left(\frac{2\underline{\theta}}{1+\underline{\theta}}, f_L\right)$. An increase in $\underline{\theta}$ increases both $\frac{2\underline{\theta}}{1+\underline{\theta}}$ and f_L . The positive effect on f_L , referred to in footnote 16, is given in the appendix and the effect on $\frac{2\underline{\theta}}{1+\underline{\theta}}$ is equal to $\frac{\partial\left(\frac{2\underline{\theta}}{1+\underline{\theta}}\right)}{\partial\underline{\theta}} = \frac{2}{(1+\underline{\theta})^2} > 0$.

segmentation occurs even when only brand-name drugs are available), the following results obtain:

Proposition 5 *The generic competition paradox can arise both when segmentation occurs before and after generic entry and when it only occurs after generic entry, although it is more likely to arise in the latter case. When segmentation occurs before and after generic entry, the paradox arises at high s , with increasingly higher s needed as f increases. When segmentation only occurs after generic entry, the paradox arises at combinations of s and f other than those involving either high s and high f or low s and low f .*

Under the assumption that segmentation occurs prior to the entry of generic drugs (that is, when p_b^C and p_b^B are compared), for a given additional coverage on the purchase of generic drugs (i.e., for a given t such that the additional coverage is $t\theta$) and a given market size (i.e., for a given $\underline{\theta}$ such that the market size is $1 - \underline{\theta}$), the higher the degree of substitutability between brand-name and generic drugs, the more likely it is for the price of brand-name drugs to increase following the introduction of generic drugs. This is consistent with the product space differentiation explanations of past work, including Perloff et al. (1995). Furthermore, as production costs increase and/or the willingness to pay for drugs decreases (i.e., $f \uparrow$ such that the profitability of the industry decreases through a closing of the gap between willingness to pay for drugs and production costs), the degree of substitutability between the two types of drugs has to be increasingly larger for the price of brand-name drugs to increase following the introduction of generic drugs. Under the assumption that segmentation does not occur prior to the entry of generic drugs (that is, when p_b^C and p_b^A are compared), for a given t and $\underline{\theta}$, the price of brand-name drugs decreases following the introduction of generic drugs at either high values of both s and f or at low values of both s and f .

The above results are depicted in Figure 4 where the zero-price differential isovalue curves are graphed in terms of s as a function of f for a given t and $\underline{\theta}$. The $p_b^C - p_b^B = 0$ curve appears in the top left corner of the Figure and defines, for $f > 0.095$ which is required for $\theta_B < 1$, the light grey area within which the price of brand-name drugs increases following the introduction of generic entry when market segmentation based on insurance coverage is practised prior to the entry of generic entry. The $p_b^C - p_b^A = 0$ curve appears in the top

right and bottom left corners of the Figure and defines the dark grey areas within which the price of brand-name drugs declines independently of whether market segmentation based on insurance coverage is practised prior to the entry of generic entry. For the bottom left dark grey area, the relevant region for an interior solution for θ_C excludes sections U_2 and U_3 : in section U_2 , no solution for θ_C exists; in section U_3 , the only valid solution is $\theta_C = 1$, that is, a corner solution.²⁰ The behavior of the price differential when the equilibrium is characterized by a corner solution (in region U_1 , or the area to the left of the vertical line at $f = 0.095$, and region U_3 as illustrated in Figure 4) is described in Figures 4a and 4b for $t = \underline{\theta} = 0.05$, with the light (dark) grey areas giving combinations of s and f for which the price of brand-name drugs increases (decreases) in response to generic entry both when market segmentation based on insurance coverage occurs prior to generic entry (in Figure 4a) and when it does not (in Figure 4b).²¹

Proposition 6 *The generic competition paradox is more likely to arise at higher t and/or lower $\underline{\theta}$.*

Under the assumption that segmentation occurs prior to the entry of generic drugs, the higher the additional coverage on the purchase of generic drugs is (i.e., as $t \uparrow$), the lower is the degree of substitutability between brand-name and generic drugs that is needed for the price of brand-name drugs to increase following the introduction of generic drugs; conversely, the smaller is the size of the market for drugs (i.e., as $\underline{\theta} \uparrow$), the higher is the degree of substitutability between the two types of drugs that is needed for the price of brand-name drugs to increase following the introduction of generic drugs. An increase in $\underline{\theta}$ also results in an increase in the lowest acceptable value of f (i.e., $\frac{2\theta}{(1+\theta)}$) so that, effectively, there exist fewer combinations of values of s and f which give rise to the GCP. Under the assumption that segmentation does not occur prior to generic entry, as t increases, the price increase

²⁰In section U_2 , the necessary (but not sufficient) condition for $\theta_C > \underline{\theta}$, that is, $D < 0$, is violated. In section U_3 , the condition for $\theta_C < 1$, that is, $-\frac{N}{D} - (1 + \underline{\theta}) < 0$, is violated. In section U_1 , which is the area to the right of $f = 0.095$, the condition for $\theta_B < 1$, that is, $f < \frac{2\theta}{(1+\theta)}$, is violated.

²¹In both cases, $\theta_B = 1$ in the shaded areas to the left of the vertical line at $f = 0.095$ while $\theta_C = 1$ in the shaded areas to the right of the vertical line at $f = 0.095$ and, for $f \leq 0.095$, in the $abde$ area and the left light grey in the bottom left corner. In the light (dark) grey areas, p_b^C is greater (lower) than p_b^B in Figure 4a and greater (lower) than p_b^A in Figure 4b. Furthermore, it can be shown that $p_b^B > p_b^A$ even when $\theta_B = 1$ (or $f < \frac{2\theta}{(1+\theta)}$) as $p_b^B - p_b^A = \alpha - \frac{\alpha}{(1+\theta)} - \frac{c}{2} = \alpha \left[\frac{2\theta}{(1+\theta)} - f \right] > 0$.

is more likely to be observed (or, put it differently, increasingly lower values of s and f are needed for the increase to occur). As $\underline{\theta}$ increases, the increase is less likely to occur (or increasingly higher (lower) values of s and lower (higher) values of f are needed for the increase to occur).

The above results are depicted in Figures 5, 6a, and 6b where the zero-price differential isovalue curves, expressed in terms of s as a function of f for given t and $\underline{\theta}$, shift in the directions of the arrows; hence, the GCP_B and the GCP_A areas (that is, the areas in which the Generic Competition Paradox occurs in relation to CASES B and A) increase in response to an increase in t (Figure 5) but decrease in response to an increase in $\underline{\theta}$ (Figures 6a and 6b). The effects of an increase in $\underline{\theta}$ are actually more complicated than above described; this is because, as $\underline{\theta}$ increases, fewer combinations of s and f yield an interior solution. Hence, with the zero-price differential isovalue curve for an interior solution (and thus for $f > 0.40$) shifting leftward, as indicated by the arrow in the top-left corner of Figure 6a, in response to an increase in $\underline{\theta}$, larger values of s are needed for given values of f in order for the price of brand-name drugs to increase following generic entry in CASE B (that is, the GCP becomes less likely to occur as reflected by the area labelled U_4). If corner solutions are however accounted for, it is quite possible for the GCP to become more likely to occur. In Figure 6a, the $p_b^C - p_b^B = 0$ curves are given for $\theta_C = \theta_B = 1$ (shaded area to the left of the $f = 0.40$ line and below the ab segment), $\underline{\theta} < \theta_B < \theta_C = 1$ (shaded area between the $f = 0.40$ line and the bde segment), and $\underline{\theta} < \theta_C < \theta_B = 1$ (shaded area to the left of the $f = 0.40$ line and above the ab segment), and accordingly labelled, and define, under the assumption that market segmentation is practised by the producer of brand-name drugs prior to generic entry, additional areas in which the GCP occurs (light-grey areas excluding the area labelled U_4 which gives combinations of s and f such that the GCP occurs and an interior solution results) and additional areas in which the GCP does not occur (dark-grey areas). In comparison to Figure 4a, which gives information about when the GCP arises with a corner solution in at least one of the two cases being compared (CASE B and/or CASE C) for $\underline{\theta} = 0.05$, the GCP results for an increasingly larger number of combinations of low values of s and f as the area in which an interior solution for both cases obtains (that is, the area to the left of the $f = 0.40$ line and below the ab segment) gets increasingly larger. In Figure 6b,

the $p_b^C - p_b^A = 0$ curves are given for $\theta_C = 1$ (area represented by $U_5 + U_6$) and for $\underline{\theta} < \theta_C < 1$, and accordingly labelled. In comparison to Figures 4a and 4c, which give information about the GCP when market segmentation is absent prior to generic entry for an interior solution and a corner solution, there are increasingly more combinations of values of s and f yielding an interior solution which do not give rise to the GCP (these additional combinations are shown in the dark-grey area which excludes U_7 and U_5 as the former applies to the $\underline{\theta} = 0.05$ case as well and the latter applies for $\theta_C = 1$). When corner solutions are considered (that is, areas U_5 and U_6), there are however additional combinations of values of s and f (that is, area U_6) at which the GCP results.

3 Numerical Example

In this section, the price of brand-name drugs is computed for each of the above cases (A through E) under the assumptions that $\alpha = 10$, which is non-consequential for the price comparisons, and that $t = \underline{\theta} = 0.05$ (that is, the best available insurance package entails a 95 percent coverage for brand-name drugs and a 97.5 percent coverage for generic drugs). Prices are reported for different values of s (that is, for different degrees of substitutability between brand-name and generic drugs) and f . There are five tables, each corresponding to a different value of f ; within each table, there are five different values of s (first column). The lower-case letters in brackets next to the values of s are included in relation to Figure 7, which reproduces the relevant regions identified in Figure 4. In each table, information is also included about the segment of the market serviced by the producer of brand-name drugs when segmentation based on insurance coverage occurs (CASE B, in the absence of generic drugs, and CASE C, in the presence of generic drugs). In CASES D and E, that is, when insurance coverage considerations are ignored, the price of brand-name drugs is not affected by the entry of generic drugs by construction (through the parameter q which captures the perceived quality differential between brand-name and generic drugs) in order for the relevance of insurance coverage heterogeneity in explaining the GCP to be fully captured. Whenever the GCP arises, p_C appears with the subscript i , where $i = A$ if the paradox occurs in relation to CASE A, $i = B$ if the paradox occurs in relation to CASE B,

and $i = AB$ if the GCP arises independently of whether segmentation occurs prior to generic entry. In some instances (those denoted with the superscript *), corner solutions prevail in CASE C and both the price and the market share of brand-name drugs, p_C and $\frac{\theta_C - \underline{\theta}}{1 - \underline{\theta}}$, are computed for $\theta_C = 1$.

Table 1: Comparison of prices and market size when $f = 0.15$								
s	p_A	p_B	p_C	$p_D = p_E$	θ_B	θ_C	$\frac{\theta_B - \underline{\theta}}{1 - \underline{\theta}}$	$\frac{\theta_C - \underline{\theta}}{1 - \underline{\theta}}$
0.15 (a)	10.27	12.63	10.53 *_A	5.75	0.79	> 1	0.78	1.00*
0.25 (b)	10.27	12.63	N/A	5.75	0.79	N/A	0.78	N/A
0.50 (d)	10.27	12.63	N/A	5.75	0.79	N/A	0.78	N/A
0.75 (e)	10.27	12.63	17.81 $_{AB}$	5.75	0.79	0.54	0.78	0.52
0.95 (g)	10.27	12.63	35.05 $_{AB}$	5.75	0.79	0.33	0.78	0.29
N/A applies as $\underline{\theta} > \theta_C < 0$.								

Table 2: Comparison of prices and market size when $f = 0.25$								
s	p_A	p_B	p_C	$p_D = p_E$	θ_B	θ_C	$\frac{\theta_B - \underline{\theta}}{1 - \underline{\theta}}$	$\frac{\theta_C - \underline{\theta}}{1 - \underline{\theta}}$
0.15 (h)	10.77	16.45	12.04 $_A$	6.25	0.61	0.86	0.59	0.85
0.25 (i)	10.77	16.45	10.73*	6.25	0.61	> 1	0.59	1.00*
0.50 (j)	10.77	16.45	10.61*	6.25	0.61	> 1	0.59	1.00*
0.75 (k)	10.77	16.45	17.87 $_{AB}$	6.25	0.61	0.52	0.59	0.49
0.95 (l)	10.77	16.45	32.37 $_{AB}$	6.25	0.61	0.33	0.59	0.29

Table 3: Comparison of prices and market size when $f = 0.50$								
s	p_A	p_B	p_C	$p_D = p_E$	θ_B	θ_C	$\frac{\theta_B - \underline{\theta}}{1 - \underline{\theta}}$	$\frac{\theta_C - \underline{\theta}}{1 - \underline{\theta}}$
0.15 (m)	12.02	23.65	20.10 $_A$	7.50	0.42	0.49	0.39	0.46
0.25 (o)	12.02	23.65	18.16 $_A$	7.50	0.42	0.53	0.39	0.51
0.50 (r)	12.02	23.65	15.44 $_A$	7.50	0.42	0.62	0.39	0.60
0.75 (u)	12.02	23.65	17.95 $_A$	7.50	0.42	0.48	0.39	0.45
0.95 (v)	12.02	23.65	25.67 $_{AB}$	7.50	0.42	0.32	0.39	0.28

Table 4: Comparison of prices and market size when $f = 0.75$								
s	p_A	p_B	p_C	$p_D = p_E$	θ_B	θ_C	$\frac{\theta_B - \underline{\theta}}{1 - \underline{\theta}}$	$\frac{\theta_C - \underline{\theta}}{1 - \underline{\theta}}$
0.15 (w)	13.27	29.33	25.74 $_A$	8.75	0.34	0.37	0.31	0.34
0.25 (y)	13.27	29.33	23.70 $_A$	8.75	0.34	0.39	0.31	0.36
0.50 (z)	13.27	29.33	19.72 $_A$	8.75	0.34	0.43	0.31	0.40
0.75 (aa)	13.27	29.33	17.89 $_A$	8.75	0.34	0.42	0.31	0.39
0.95 (ab)	13.27	29.33	18.96 $_A$	8.75	0.34	0.32	0.31	0.28

Table 5: Comparison of prices and market size when $f = 0.95$								
s	p_A	p_B	p_C	$p_D = p_E$	θ_B	θ_C	$\frac{\theta_B - \theta}{1 - \theta}$	$\frac{\theta_C - \theta}{1 - \theta}$
0.15 (<i>ad</i>)	14.27	33.29	29.49 _A	9.75	0.30	0.31	0.26	0.27
0.25 (<i>ae</i>)	14.27	33.29	27.23 _A	9.75	0.30	0.32	0.26	0.28
0.50 (<i>ag</i>)	14.27	33.29	22.24 _A	9.75	0.30	0.35	0.26	0.32
0.75 (<i>ah</i>)	14.27	33.29	17.64 _A	9.75	0.30	0.36	0.26	0.33
0.95 (<i>ai</i>)	14.27	33.29	13.59	9.75	0.30	0.30	0.26	0.26

4 Concluding Remarks

In this paper, the relevance of insurance coverage heterogeneity in explaining the generic competition paradox, that is, the observation that the price of brand-name drugs increases following the introduction of generic drugs, is examined in a two-stage game involving a single producer of brand-name drugs and n quantity-competing producers of generic drugs.

Past theoretical work attempting to explain the paradox has relied on product differentiation, exogenous market segmentation, and brand loyalty. Here we suggest that, even in cases where differentiation and brand loyalty alone would not result in brand name price increases, endogenous segmentation on the basis of insurance coverage by brand name producers can reverse the decline in prices caused by generic competition. While empirical work has indicated that insurance may play an important role in the paradox, no theoretical study to date has examined the role of endogenous segmentation by insurance coverage.

In the model presented, when consumers are homogenous in their insurance coverage, the generic competition paradox can arise provided that consumers' perceived quality differential between brand-name and generic drugs exceeds some threshold level which depends positively upon the degree of substitutability between the two types of drugs and the willingness to pay for drugs, and negatively upon marginal production costs. With the perceived quality differential at a level which does not give rise to price increases following the introduction of generic drugs when consumers have similar coverages (that is, with the brand loyalty argument assumed away), heterogeneity in insurance coverage is shown to lead to the paradox not only when market segmentation based on insurance coverage is absent in the pre-entry stage, so that a price increase can be expected as a result of the market shrinkage due to the segmentation, but also when market segmentation occurs prior to generic entry, although the

paradox is less likely to arise in the latter case. In particular, and independently of whether market segmentation is employed prior to generic entry, a price increase for brand-name drugs is likely to result as generic drugs become available at combinations of high values of s and low values of f , that is, when the two types of drugs are highly substitutable and the market is profitable. When market segmentation does not occur prior to generic entry, the paradox is also likely to arise at combinations of low values of s and high values of f , that is, when the two types of drugs are highly differentiated and the market is not profitable. While the extent of substitutability between the two types of drugs is an important factor in the determination of how the price of brand-name drugs adjusts as generic drugs make their way into the pharmaceutical market, although its effects runs counterintuitive with respect to previous studies (e.g., Perloff et al., 1995), market profitability as captured by f is equally important.

Within the context of our model with only interior solutions being considered, it is then possible for the paradox to arise when the two types of drugs are highly substitutable in some cases (consistent with other studies which rely on product differentiation) but not in other cases, depending on how profitable the industry is. Furthermore, under the assumption that market segmentation occurs only after generic entry, the paradox may arise when the two types of drugs are either highly substitutable (and the industry highly profitable) or highly differentiated (and the industry highly unprofitable). The set of feasible combinations of values of s and f (with s measuring product substitutability and f industry profitability) at which the paradox is observable is also dependent upon the size of the market (θ) and the additional coverage on the purchase of generic drugs for a given insurance package (t). Specifically, as the market expands (this is equivalent to an improvement in the best insurance package available) and/or insurance companies decrease deductables applied on the purchase of generic drugs, the Generic Competition Paradox becomes more likely to arise.

In essence, at low levels of product differentiation, the producer of brand-name drugs responds to generic entry by providing its product to fewer consumers (those with better insurance coverage) and can thus supply it at a higher price; as the market becomes less profitable and/or the best insurance package involves a lower coverage and/or the difference in deductables between the two types of drugs decreases, the incentive to shrink the segment

of the market buying brand-name drugs weakens making the GCP less likely to occur. On the other hand, at high levels of product differentiation, the producer of brand-name drugs responds to generic entry by offering its product to more consumers, and must thus charge a lower price, when it practises market segmentation prior to generic entry but to fewer consumers, and can thus charge a higher price, when it does not practise market segmentation prior to generic entry. In the latter case, as the market shrinks in terms of insurance packages available (with the best coverage decreasing) and/or insurance companies offer a lower deductible on the purchase of generic drugs, the incentive to supply to fewer consumers weakens so that the GCP becomes less likely to occur or, equivalently, the GCP is more likely to arise when the market is not profitable.

References

- [1] Anis, A.H. (1992), “Pharmaceutical Prices with Insurance Coverage and Formularies,” *Canadian Journal of Economics*, 25(2): 420-437.
- [2] Berndt, E.R., Z. Griliches, and J.G. Rosett (1993), “Auditing the Producer Price Index: Micro-Evidence from Prescription Pharmaceutical Preparations,” *Journal of Business and Economic Statistics* 11(July), 251-64.
- [3] Bhattacharya, J., and W.B. Vogt (2003), “A Simple Model of Pharmaceutical Price Dynamics,” *Journal of Law and Economics*, 46(October), 599-626.
- [4] Brekke, K.R., A.L. Grasdal, and T.H. Holmås (2009), “Regulation and pricing of pharmaceuticals: Reference pricing or price cap regulation?” *European Economic Review* 53(2), 170-185.
- [5] Caminal, R., and X. Vives (1999), “Price Dynamics and Consumer Learning,” *Journal of Economics and Management Strategy*, 8(1): 95-131.
- [6] Caves, R.E., M.D. Whinston, and M.A. Hurwitz (1991), “Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry,” *Brookings Papers on Economic Activity (Microeconomics)*, 1-4.8
- [7] Davis, S.J., K.M. Murphy, and R.H. Topel (2004), “Entry, Pricing, and Product Design in an Initially Monopolized Market,” *Journal of Political Economy*, 112(1): S188-S225.
- [8] Danzon, P.M., and L-W. Chao (2000), “Does Regulation Drive Out Competition in Pharmaceutical Markets?” *Journal of Law and Economics*, 43(October): 311-357.
- [9] Dietz, E. (2004), “Trends in Employer-Provided Prescription-Drug Coverage,” *Monthly Labour Review Online*, 127(8): 37-45.
- [10] Ferrara, I., and Y. Kong (2008), “Can Health Insurance Coverage Explain the Generic Competition Paradox?,” *Economics Letters*, 101, 48-52.

- [11] Frank, R.G., and D.S. Salkever (1992), “Pricing Patent Loss and the Market for Pharmaceuticals,” *Southern Economic Journal*, 59(2): 165-179.
- [12] Frank, R.G., and D.S. Salkever (1997), “Generic Entry and the Pricing of Pharmaceuticals,” *Journal of Economics and Management Strategy*, 6(1): 75-90.
- [13] Grabowski, H.G., and J.M. Vernon (1992), “Brand Loyalty, Entry and Price Competition in Pharmaceuticals After the 1984 Drug Act,” *Journal of Law and Economics*, 35(2): 331-350.
- [14] Griliches, Z., and I.M. Cockburn (1994), “Generics and New Goods in Pharmaceutical Price Indexes,” *American Economic Review* 84(December), 1213-1232.
- [15] Hellerstein, J. K. (1994), “The Demand for Post-Patent Prescription Pharmaceuticals,” *RAND Journal of Economics*, 29(1): 108-136.
- [16] Hurwitz, M.A., and R.E. Caves (1988), “Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals,” *Journal of Law and Economics*, 31(2): 299-320.
- [17] Kamien, M.I., and I. Zang (1999), “Virtual Patent Extension by Cannibalization,” *Southern Economic Journal*, 66(1), 117-131.
- [18] Kong, Y. (2004), “The Price Premium of Generic Pharmaceuticals to Brand-names and Pharmaceutical Price Index,” *Applied Economics*, 36(7): 731-737.
- [19] OECD (2007), *OECD Health Data 2007: Statistics and Indicators for 30 Countries*, OECD, Paris.
- [20] Pavcnik, N. (2002), “Do Pharmaceutical Prices Respond to Potential Patient Out-Of-Pocket Expenses?,” *RAND Journal of Economics*, 33(3): 469-487.
- [21] Perloff, J.M., V.Y. Suslow, and P.J. Seguin (1995), “Higher Prices from Entry: Pricing of Brand-Name Drugs,” *mimeo*.

- [22] Salop, S.C. (1979), “Monopolistic Competition with Outside Goods,” *Bell Journal of Economics and Management Science* 10(Spring), 141-156.
- [23] Scherer, F. M. (1993), “Pricing, Profits, and Technological Progress in the Pharmaceutical Industry,” *Journal of Economic Perspectives*, 7(3): 95-115.
- [24] Wiggins, S.N., and R. Maness (2004), “Price Competition in Pharmaceutical Markets: The Case of Anti-infectives,” *Economic Inquiry*, 42(2): 247-263.

5 APPENDIX

INTERIOR SOLUTION OF θ_C : Conditions

From (39), an interior solution of θ_C , that is, $\underline{\theta} < \theta_C < 1$, requires that $2\underline{\theta}^2 < -\frac{N}{D} < (1 + \underline{\theta})$, where N and D are given in (40) and (41). Specifically, $D < 0$ for $\theta_C > 0$, $-\frac{N}{D} - 2\underline{\theta}^2 > 0$ for $\theta_C > \underline{\theta}$, and $-\frac{N}{D} - (1 + \underline{\theta}) < 0$ for $\theta_C < 1$. With f_L and f_H denoting values of f solving $-\frac{N}{D} = (1 + \underline{\theta})$ and $-\frac{N}{D} = 2\underline{\theta}^2$, namely,

$$f_L = \frac{(2-s)(1-s)(4\underline{\theta}+2s) + s(4\underline{\theta}-2s)}{s(4\underline{\theta}-2s) - (1+\underline{\theta})(2-s)[(1-t)s - (2-s^2)]}$$

and

$$f_H = \frac{-2(1+\underline{\theta}+\underline{\theta}s)(2-s)(1-s) - 2(1+\underline{\theta})s + 2\underline{\theta}s^2}{\underline{\theta}(1+\underline{\theta})(2-s)[(1-t)s - (2-s^2)] - 2(1+\underline{\theta})s + 2\underline{\theta}s^2},$$

the following obtains: (i) for $f < f_L < f_H$, $-\frac{N}{D} < 2\underline{\theta}^2 < (1 + \underline{\theta})$; (ii) for $f_L < f < f_H$, $2\underline{\theta}^2 < -\frac{N}{D} < (1 + \underline{\theta})$; (iii) for $f > f_H > f_L$, $-\frac{N}{D} > (1 + \underline{\theta}) > 2\underline{\theta}^2$. Hence, when CASES B and C are compared, that is, when segmentation based on insurance coverage occurs both before and after generic entry, the lower-bound value of f (f_{\min}) is given by $f_{\min} = \max\left(\frac{2\underline{\theta}}{1+\underline{\theta}}, f_L\right)$, with $f > \frac{2\underline{\theta}}{1+\underline{\theta}}$ required for an interior solution of θ_B , while the upper-bound value (f_{\max}) is given by $f_{\max} = \min(f_H, 1)$. When CASES A and C are compared, that is, when segmentation occurs only after generic entry, $f_{\min} = f_L$ and $f_{\max} = \min(f_H, 1)$. It turns out that, unless t and $\underline{\theta}$ are excessively high, $f_H > 1$ over the entire range of values of s so that $f_{\max} = 1$.

For given t and $\underline{\theta}$, the above f_L expression can be solved for s as a function of f such that $\theta_C = 1$. Above (or to the right of) the resulting isovalue curve in (f, s) space, $\theta_C < 1$.

As

$$\frac{\partial f_L}{\partial \underline{\theta}} = -\frac{(s-2)[(-s^4 - 6s^2 + 6s^3 + 4s)t + (s-1)(s^2-2)(s-2)^2]}{(2\underline{\theta}st - \underline{\theta}s^2t + \underline{\theta}s^3 - \underline{\theta}s^2 + 4\underline{\theta}a - 4s + 2st + 4 - 3s^2 - s^2t + s^3)^2} > 0$$

and

$$\frac{\partial f_L}{\partial t} = \frac{2(4\underline{\theta} + 2\underline{\theta}s^2 - 4\underline{\theta}s + s^3 + 2s - 4s^2)s(s-2)(1+\underline{\theta})}{(2\underline{\theta}st - \underline{\theta}s^2t + \underline{\theta}s^3 - \underline{\theta}s^2 + 4\underline{\theta}a - 4s + 2st + 4 - 3s^2 - s^2t + s^3)^2} < 0,$$

an increase in $\underline{\theta}$ (t) results in a rightward (leftward) shifting of the $\theta_C = 1$, thus making an interior solution less (more) likely to attain.

PROPOSITION 3: Sign of $\frac{\partial \theta_C}{\partial \underline{\theta}}$

From (40),

$$\frac{N}{2(1+\underline{\theta})} > 2\underline{\theta}s[(2-s)(1-s) - s(1-f)]$$

so that, if $(1+\underline{\theta})\theta_C^2 > \frac{N}{2(1+\underline{\theta})}\frac{N}{D^2}$, then $(1+\underline{\theta})\theta_C^2 > 2\underline{\theta}s[(2-s)(1-s) - s(1-f)]\frac{N}{D^2}$ and $\frac{\partial\theta_C}{\partial\underline{\theta}} > 0$. Upon manipulation of (39),

$$2\theta_C + \underline{\theta} = 2\sqrt{\frac{\theta^2}{4} - \frac{N}{D}} \implies \theta_C^2 + \underline{\theta}\theta_C = -\frac{N}{D} \implies \theta_C^2(\theta_C + \underline{\theta})^2 = \frac{N^2}{D^2},$$

and

$$(1+\underline{\theta})\theta_C^2 - \frac{N}{2(1+\underline{\theta})}\frac{N}{D^2} = \frac{1}{2(1+\underline{\theta})}[(1+\underline{\theta})^2\theta_C^2 - (\theta_C + \underline{\theta})^2\theta_C^2],$$

which is positive as, for an interior solution, $\theta_C < 1$.

PROPOSITION 3: Sign of $\frac{\partial\theta_C}{\partial f}$

By (44), the sign of $\frac{\partial\theta_C}{\partial f}$ is equal to the sign of

$$A1 = (1+\underline{\theta})(2-2s+s^2)[(1-t)s - (2-s^2)] + 4s^2(1-s).$$

As $[(1-t)s - (2-s^2)] < 0$, $A1$ is maximized when $\underline{\theta} = t = 0$. For $\underline{\theta} = t = 0$,

$$A1 = (2-2s+s^2)(s-2+s^2) + 4s^2(1-s),$$

which is negative for $s < 0.8536345110$ and positive for $0.8536345110 < s < 1$. For $0.8536345110 < s < 1$, $\frac{\partial\theta_C}{\partial f} < 0$ if $ts(1+\underline{\theta}) - \underline{\theta}(s^2+s-2) > \frac{(1-s)(-s^3+4s^2+2s-4)}{s^2-2s+2}$. The inequality is least likely to hold when its left-hand side is minimized (that is, when $\underline{\theta} = 0$) and its right-hand side is maximized (that is, when $s = 0.9248386528$). For $\underline{\theta} = 0$ and $s = 0.9248386528$, the inequality is always satisfied provided that $t > 0.03878582123$.

PROPOSITION 3: Isovalue curve $t(\underline{\theta}, f, s) = k$

With $t(\underline{\theta}, f, s)$ defining the solution to $\Phi(\underline{\theta}, f, s, t) = 0$, which is explicitly given in (46), the slope of the isovalue curve $t(\cdot) = k$, for $0 \leq k \leq 1$, can be expressed as

$$\frac{ds}{df} = -\frac{t_f(\cdot)}{t_s(\cdot)},$$

where

$$\begin{aligned}
t_f(\cdot) = & \frac{(-f^2 - \underline{\theta}f^2 + 4f)s^6 + (4\underline{\theta}f^2 - 16f - 8)s^5 + (-4\underline{\theta}f^2 + 48f + 40)s^4}{f^2(1 + \underline{\theta})(fs^2 - 4s + 4)^2} + \\
& + \frac{(64f + 128)s^3 + (12f^2 + 4\underline{\theta}f^2 + 16f + 224)s^2 - (16f^2 + 16\underline{\theta}f^2 + 160)s}{f^2(1 + \underline{\theta})(fs^2 - 4s + 4)^2} + \\
& + \frac{16f^2 + 16\underline{\theta}f^2 + 32}{f^2(1 + \underline{\theta})(fs^2 - 4s + 4)^2}
\end{aligned}$$

and $t_s(\cdot)$ is given in (47). As $t_s(\cdot)$ is always negative and $t_f(\cdot) > 0$ for $f < \tilde{f}$, where

$$\begin{aligned}
\tilde{f} = & -\frac{2\sqrt{(s^4 - 4)(s^4 - 4s^3 + 12s^2 - 16s + 4)(s - 2)^2[s^2 + 2(1 + \underline{\theta})(1 - s)]}}{(1 + \underline{\theta})(s^6 + 16s - 16) + 4s^2[\underline{\theta}(-s^3 + s^2 - 1) - 3]} + \\
& + \frac{2s^2(s^4 - 4s^3 + 12s^2 - 16s + 4)}{(1 + \underline{\theta})(s^6 + 16s - 16) + 4s^2[\underline{\theta}(-s^3 + s^2 - 1) - 3]},
\end{aligned}$$

it obtains that $\frac{ds}{df}$ is positive for $f \in (0, \tilde{f})$ and negative for $f \in (\tilde{f}, 1)$. Furthermore, for $0 < f < 1 - k$, the isovalue curve $t(\cdot) = k$ lies above the $f = \hat{f}$ curve, where

$$\hat{f} = \frac{-s^2 + s + 2 - \sqrt{2s^4 - 6s^3 + 9s^2 - 12s + 8}}{s},$$

which gives combinations of f and s such that $\left.\frac{\partial t}{\partial \underline{\theta}}\right|_{\Phi=0} = 0$, as per (49). Hence, as $\left.\frac{\partial t}{\partial \underline{\theta}}\right|_{\Phi=0} < 0$ (> 0) for $f < \hat{f}$ ($f > \hat{f}$) and $t_s(\cdot) < 0$, an increase in $\underline{\theta}$ triggers an upward (downward) shifting of the $t(\cdot) = k$ isovalue curve for $f < \hat{f}$ ($f > \hat{f}$). With $s\left(f|_{\Phi=0, t=k}\right)$ denoting the s as a function of f that solves $t(\cdot) = k$ and $s\left(f|_{f=\hat{f}}\right)$ the s as a function of f that solves $f = \hat{f}$, it obtains that $s\left(f|_{\Phi=0, t=k}\right) = s\left(f|_{f=\hat{f}}\right)$ when $f = 0$ or $f = 1 - k$; by concavity of the former and convexity of the latter,²² it follows that $s\left(f|_{\Phi=0, t=k}\right) > s\left(f|_{f=\hat{f}}\right)$ for $0 < f < 1 - k$, and $s\left(f|_{\Phi=0, t=k}\right) < s\left(f|_{f=\hat{f}}\right)$ for $1 - k < f < 1$.

PROPOSITION 4: Slope of s^*

With $t^*(\theta, f, s^*)$ defining the solution to $M(\theta, f, s^*, t) = 0$, such that $\theta_C = \theta_B$, the slope of s^* , where

$$s^* = \frac{R + f[(1 + \underline{\theta})(3 + t) + 2] - 8}{2[f(1 + \underline{\theta}) - 2]},$$

as per (50), with

$$R = \sqrt{\{8 - f[(1 + \underline{\theta})(3 + t) + 2]\}^2 - 8[f(1 + \underline{\theta}) - 2][f(1 + \underline{\theta})(f + t) - 2]} > 0,$$

²²It can easily be shown that $\frac{\partial^2 f}{\partial s^2} < 0$ which implies that $s\left(f|_{f=\hat{f}}\right)$ is convex.

can be expressed as

$$\frac{ds^*}{df} = -\frac{t_f^*(\cdot)}{t_s^*(\cdot)} = -\frac{2(s-2)[s^2-4s+f^2(1+\underline{\theta})+2]}{f(s^2-4s+6-2f)[f(1+\underline{\theta})-2]},$$

where

$$t_f^*(\cdot) = \frac{2[s^2-4s+f^2(1+\underline{\theta})+2]}{(s-2)f^2(1+\underline{\theta})}$$

and

$$t_s^*(\cdot) = \frac{(s^2-4s+6-2f)[f(1+\underline{\theta})-2]}{(s-2)^2 f(1+\underline{\theta})}.$$

For $s = 2 - \sqrt{2 - f^2(1 + \underline{\theta})}$, $\frac{ds^*}{df} = 0$ and $\frac{d^2s^*}{df^2} = -\frac{4(1+\underline{\theta})\sqrt{2-f^2(1+\underline{\theta})}}{[f(1+\underline{\theta})-2][f^2(1+\underline{\theta})+2f-4]} < 0$, so that s^* is increasing for $f \in \left(0, \sqrt{\frac{-s^2+4s-2}{1+\underline{\theta}}}\right)$ and decreasing for $f \in \left(\sqrt{\frac{-s^2+4s-2}{1+\underline{\theta}}}, 1\right)$.